

IN THE CLAIMS:

1. **(Previously Amended)** A system comprising:
a decoder responsive to an input signal stream comprising text commingled with FAP information, that separates the FAP information from the text, and develops phonemes from said text,
a converter responsive to said decoder, that converts said phonemes to additional FAP information and outputs said additional FAP information combined with said FAP information separated by said decoder, and
a face rendering module responsive to an applied face model signal and to said output developed by said converter.

C/ 2. **(Previously Amended)** A method for transmitting signals to apparatus that produces sounds and includes a video synthesizer comprising the steps of:
generating a first signal stream that includes signals for generating said sounds;
generating a second signal stream of commands to said video synthesizer, which commands comprise FAP information that excludes viseme information; and
combining said first signal stream with said second signal stream to form a signal stream for said transmitting.

3. **(Deleted)** .

4. **(Deleted)**.

5. **(Deleted)**.

6. **((Deleted))**.

7. **((Deleted))**.

8. **((Deleted))**.

9. (Deleted).

10. (Deleted).

11. (Deleted).

12. (Previously Amended) Apparatus comprising:

a decoder, responsive to an input signal comprising signals representing audio and embedded video synthesis command signals, that separates the command signals from the signals representing audio to develop an audio signal stream and a video synthesis command signals stream,

a converter responsive to said audio signal stream for developing sound, and

a video synthesizer responsive to said video synthesis command signals stream for developing images.

13. (Previously Added) The apparatus of claim 12 where said signals representing audio comprise text, and said converter is a speech synthesizer responsive to said text.

14. (Previously Added) The apparatus of claim 12 where said signals representing audio comprise text, said decoder, following separation of said command signals from said input signal, converts said text to elemental sound element signals and applies said sound element signals to said converter, and said converter is an audio synthesizer that is adapted to respond to said sound element signals.

15. (Previously Added) The apparatus of claim 14 where said converter is a speech synthesizer.

16. (Previously Added) The apparatus of claim 12 where

said audio signal stream comprises text,
said decoder, following separation of said command signals from said input
signal, converts said text to phoneme signals, and
said converter is a speech synthesizer responsive to said phoneme signals.

17. (Previously Added) The apparatus of claim 16 where said video synthesis
command signals are FAPs.

18. (Previously Added) The apparatus of claim 17 where said video synthesizer
includes an input for receiving synthesis parameters.

19. (Previously Added) The apparatus of claim 18 where said synthesis
parameters are face model parameters.

20. (Currently Amended) Apparatus comprising:
a decoder, responsive to an input signal comprising signals representing audio and
embedded video synthesis command signals, that separates the command signals from the
signals representing audio, and converts text signals found in said audio into phoneme
signals, to develop thereby an audio signal stream and a first set of video synthesis
command signals stream;

~~The apparatus of claim 16 further comprising a converter for generating a second set of~~
~~additional video synthesis command signals, over and above said video synthesis~~
~~command signals stream, from said phoneme signals; and applying said additional video~~
~~synthesis command signals generated by said converter to said video synthesizer, in~~
~~addition to said video synthesis command signals stream being applied to said video~~
~~synthesizer~~

a speech synthesizer responsive to said audio signal stream for developing sound;
and

a video synthesizer responsive to said video synthesis command signals stream
for developing images.

21. (Previously Added) The apparatus of claim **20** where said converter is interposed between said decoder and said video synthesizer, merging said command signals separated from said input signal with said command signals generated by said converter, to form a single stream of input-signal-related command signals that is applied to said video synthesizer.

22. (Previously Added) The apparatus of claim **21** where said converter generates additional command signals interposed between said input-signal-related command signals.

23. (Previously Added) The apparatus of claim **20** where said video synthesis command signals are FAPs, and said video synthesis command signals generated by said converter are FAPs.

24. (Previously Added) The apparatus of claim **23** where said video synthesis command signals generated by said converter are members of the FAP 1 parameter.

25. (Previously Added) The apparatus of claim **23** where said video synthesis command signals generated by said converter are members of the FAP 1 parameter or FAP3-68 parameters, inclusively.

26. (Previously Added) The apparatus of claim **12** where said decoder generates additional command signals that interpolate between the separated command signals from said input signal.

27. (Previously Added) The apparatus of claims **26** or **22** where each set of said additional command signals that are interposed between a pair of command signals interpolates between said pair of command signals.

28. (Previously Added) The apparatus of claim **27** where said video synthesizer generates images at a selected frame rate, and said interpolation generates a command signal for each frame.

29. (Previously Added) The apparatus of claim **27** where said interpolation follows a function that is of an order higher than 2.

30. (Previously Added) The apparatus of claim **27** where said interpolation follows a function that is of order 4.

31. (Previously Amended) A method comprising the steps of:
receiving an input signal that comprises signals representing audio and embedded video synthesis command signals;
separating said input signal into an audio signal stream and a video synthesis command signals stream;
converting said audio signal stream to audio, and
synthesizing at least one image from said video synthesis command signals stream with aid of a FAP-based face model.

32. (Previously Added) The method of claim **31** where said signals representing audio comprise text, and said step of converting synthesizes speech.

33. (Previously Added) The method of claim **32** further comprising a step of converting said text into phonemes, and said step of converting synthesizes speech from said phonemes.

34. (Previously Added) The method of claim **31** where said video synthesis command signals comprise Facial Animation Parameter signals.

35. (Previously Deleted).

36. (Previously Added) The method of claim **33** further comprising a step of generating video synthesis command signals from said phonemes and said step of synthesizing is responsive to a combined command signals stream that includes said command signals developed in said step of separating and said command signals generated in said step of generating.

37. (Previously Added) The method of claim **36** where said command signals comprise Facial Animation Parameter signals.

38. (Previously Added) The method of claim **36** further comprising a step of developing a plurality of additional command signals interposed between command signals of said combined command signals stream.

39. (Previously Added) The method of claim **38** where said step of synthesizing generates images at a selected frame rate, and said step of developing develops said additional command signals to provide a command signal for each frame.

40. (Previously Added) The method of claim **38** where said step of developing develops a set of said additional command signals between each pair of said command signals of said combined command signals stream, and said set of additional command signals interpolated between said pair of said command signals of said combined command signals stream.

41. (Currently Amended) ~~The method of claim 40~~ A method comprising the steps of:

receiving an input signal that comprises text signals representing audio and embedded video synthesis command signals;

separating said input signal into an audio signal stream and a video synthesis command signals stream;

converting said text signal stream into phonemes and synthesizing speech from said phonemes, and

developing a plurality of additional command signals and interposing the additional command signals into said video synthesis command signals stream to form a combined command signals stream; and

synthesizing at least one image from said video synthesis command signals stream with aid of a FAP-based face model;

where said step of developing develops a set of said additional command signals between each pair of said command signals of said combined command signals stream, and said set of additional command signals interpolated between said pair of said command signals of said combined command signals stream; and

where said interpolation is in accord with a function of order greater than 2.

42. (Currently Amended) The method of claim ~~40~~ **41** where said interpolation is in accord with a function of order 4.

43. (Previously Amended) Apparatus comprising

A decoder/synthesize module that is responsive to an input stream that includes a text specification commingled with explicit FAP information, outputting a synthesized voice at a first output, and phonemes as well as said FAP information at a second output;

a converter responsive to said second output for generating a sequence of facial animation parameters;

face rendering module responsive to said converter; and

a compositor, responsive to said synthesizer and to said face rendering module.

44. (Previously Added) The apparatus of claim 43, further adapted to accept said input from a remote location that is communicated to said apparatus via a communication network.

45. (Previously Added) The apparatus of claim 43 where said FAP information that is explicitly included in said input comprises interspersed bookmarks.

46. (Previously Added) The apparatus of claim 45 where each bookmark conveys information about identity of a FAP, and ultimate state of the FAP.

47. (Previously Added) The apparatus of claim 46 where said information conveys amplitude information

48. (Previously Added) The apparatus of claim 46 where said information conveys a duration measure for transiting to specified state.

49. (Previously Added) The apparatus of claim 46 where the said ultimate state of the FAP is reached in accordance with a specified transition path.

50. (Previously Added) The apparatus of claim 49 where the transition path is selected by said facial animation module.

51. (Previously Added) The apparatus of claim 49 where said transition path is specified by the bookmark.

52. (Currently Amended) ~~The apparatus of claim 49~~ Apparatus comprising
A decoder/synthesize module that is responsive to an input stream that includes a
text specification commingled with explicit FAP information, in the form of interspersed
bookmarks, each conveying information about identity of a FAP and an ultimate state
that the FAP reaches in accordance with a specified transition path, outputting a
synthesized voice at a first output, and phonemes as well as said FAP information at a
second output;

a converter responsive to said second output for generating a sequence of facial
animation parameters;

face rendering module responsive to said converter; and

a compositor, responsive to said synthesizer and to said face rendering module;

where the transition path follows the equation $f(t) = a_s + (a - a_s)t$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, and t is

time, ranging between 0 and 1, or a transition path that involves higher powers of t or e raised to power t .

53. (Currently Amended) The apparatus of claim 49 52 where the transition path follows the equation $f(t) = a_s + (1 - e^{-t})(a - a_s)$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, and t is time, ranging between 0 and 1.

54. (Currently Amended) The apparatus of claim 49 52 where the transition path follows the equation $f(t) = a_s + \frac{(a - a_s)}{(1 - e^{-\lambda(t - FABdur/2)})}$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, $FABdur$ is specified in said bookmark, λ is a specified parameter, and t is time, ranging between 0 and 1.

55. (Currently Amended) The apparatus of claim 49 52 where the transition path follows the equation $f(t) = a_s + (2t^3 - 3t^2 + 1) + (-2t^3 + 3t^2)a + (t^3 - 2t^2 + t)g_s$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, g_s is a specified parameter, and t is time, ranging between 0 and 1.

56. (Currently Amended) The apparatus of claim 49 52 where the FAP amplitude transition path follows the equation $FAP_{amp}(t) = startVal_i(2t^3 - 3t^2 + 1) + FAP_{val}_i(-2t^3 + 3t^2) + startTan_i(t^3 - 2t^2 + 1)$, where $startVal$, FAP_{val} , and $startTan$, are specified constants.

57. (Previously Amended) A method comprising the steps of:
receiving an input that includes a text specification commingled with explicit FAP information, and outputting a synthesized voice at a first output, and phonemes as well as said FAP information at a second output;
generating a sequence of facial animation parameters from signals of said second output;

rendering images from output signals developed by said step of generating; and combining said synthesized voice and said images.

58. (Previously Added) The method of claim 57, where said step of receiving accepts said input from a remote location that is communicated to said apparatus via a communication network.

59. (Previously Added) The method of claim 57 where said FAP information that is explicitly included in said input comprises interspersed bookmarks.

60. (Previously Added) The method of claim 59 where each bookmark conveys information about identity of a FAP, and ultimate state of the FAP.

61. (Previously Added) The method of claim 60 where said information conveys amplitude information

62. (Previously Added) The method of claim 60 where said information conveys a duration measure for transiting to specified state.

63. (Previously Added) The method of claim 60 where the said ultimate state of the FAP is reached in accordance with a specified transition path.

64. (Previously Added) The method of claim 63 where the transition path is selected by said facial animation module.

65. (Previously Added) The method of claim 63 where said transition path is specified by the bookmark.

66. (Currently Amended) ~~The method of claim 63~~ A method comprising the steps of:

receiving an input that includes a text specification commingled with explicit FAP information in the form of interspersed bookmarks, each conveying information about identity of a FAP and an ultimate state that the FAP reaches in accordance with a specified transition path, outputting a synthesized voice at a first output, and phonemes as well as said FAP information at a second output

generating a sequence of facial animation parameters from signals of said second output;

rendering images from output signals developed by said step of generating; and combining said synthesized voice and said images;

where the transition path follows the equation $f(t) = a_s + (a - a_s)t$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, and t is time, ranging between 0 and 1, or a transition path that involves higher powers of t or e raised to power t .

67. (Currently Amended) The method of claim ~~63~~ 66 where the transition path follows the equation $f(t) = a_s + (1 - e^{-t})(a - a_s)$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, and t is time, ranging between 0 and 1.

68. (Currently Amended) The method of claim ~~63~~ 66 where the transition path follows the equation $f(t) = a_s + \frac{(a - a_s)}{(1 - e^{-\lambda(t - FABdur/2)})}$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, $FABdur$ is specified in said bookmark, λ is a specified parameter, and t is time, ranging between 0 and 1.

69. (Currently Amended) The method of claim ~~49~~ 66 where the transition path follows the equation $f(t) = a_s + (2t^3 - 3t^2 + 1) + (-2t^3 + 3t^2)a + (t^3 - 2t^2 + t)g_s$, where a_s is amplitude measure at beginning of transition, a is specified in said bookmark, g_s is a specified parameter, and t is time, ranging between 0 and 1.

70. (Currently Amended) The method of claim ~~63~~ 66 where the FAP amplitude transition path follows the equation

$$FAP_{amp}(t) = startVal_i(2t^3 - 3t^2 + 1) + FAP_{val}_i(-2t^3 + 3t^2) + startTan_i(t^3 - 2t^2 + 1), \text{ where}$$

$startVal$, FAP_{val} , and $startTan$, are specified constants.